FIDeS: Frühwarn- und Intrusion Detection-System auf der Basis von kombinierten Methoden der KI

Workshop Informationssicherheit

Carsten Elfers

06.11.2009

Rahmenbedingungen

- Fördermittelgeber: Bundesministerium für Bildung und Forschung (BMBF)
- 6 Mitarbeiter TZI, insg. 14 Mitarbeiter
- Projektlaufzeit: 01.09.2008 – 31.08.2011
- Projekt-Konsortium
  - TZI/Universität Bremen
  - ZF Friedrichshafen AG
  - nicos AG
  - algorithmica technologies GmbH
  - if(is)/Fachhochschule Gelsenkirchen
  - T-Systems Enterprise Services GmbH
  - mobile solution group GmbH

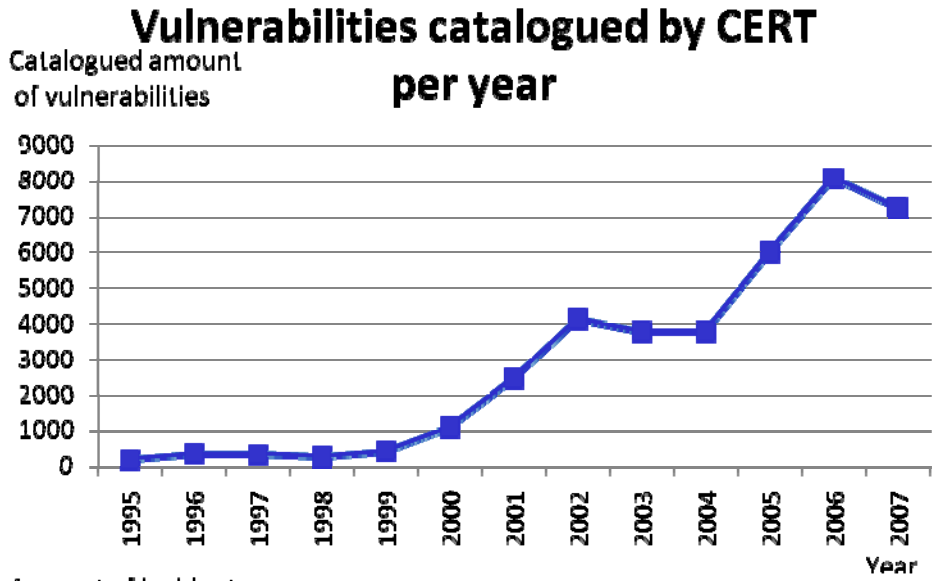
➔ Guter Mix aus Industrie- und Forschungspartnern

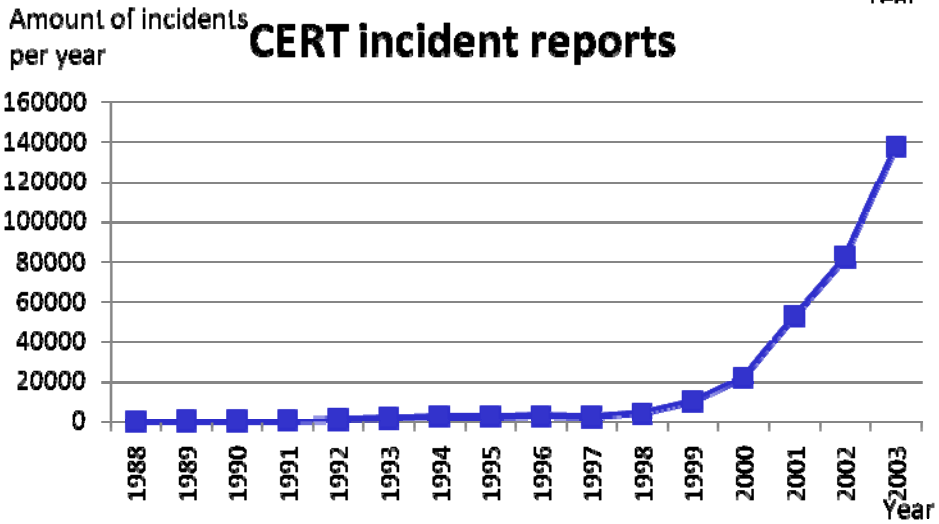
Überblick

1. Motivation, Ziele, Architektur
2. FIDeS Komponenten
   - Vorverarbeitung
   - Wissensbasis
   - Logische Repräsentation
   - Wahrscheinlichkeitstheoretische Repräsentation
   - Benutzerschnittstelle
3. Zusammenfassung

Einleitung

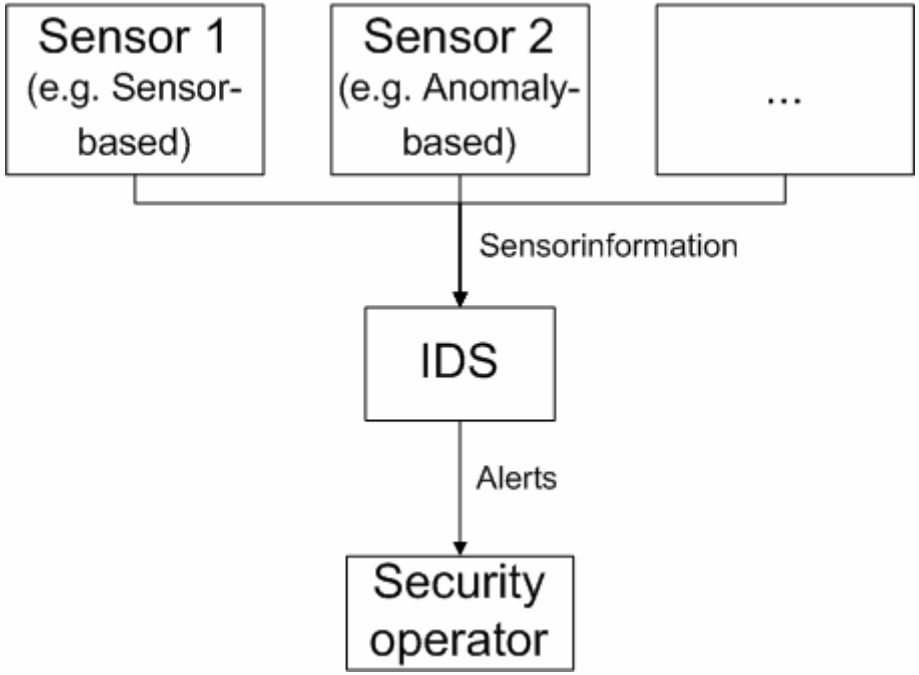
- Angriffsanzahl und Vielzahl steigt stetig
- Angriffserkennung ist notwendig um Sicherheit zu steigern
- Expertenwissen ist erforderlich um (tatsächliche) Angriffe zu identifizieren





Grundlegende Verfahren

- Anomalie-basierte Sensoren
  - Hohe falsch positiv Rate
- Signatur-basierte Sensoren
  - Erkennt nur bekannte Angriffe
- Sensorinformationen werden in Intrusion Detection Systemen gesammelt
- Korrelation erfolgt nur rudimentär automatisch oder manuell durch Experten



Auswirkungen von Fehleinschätzungen

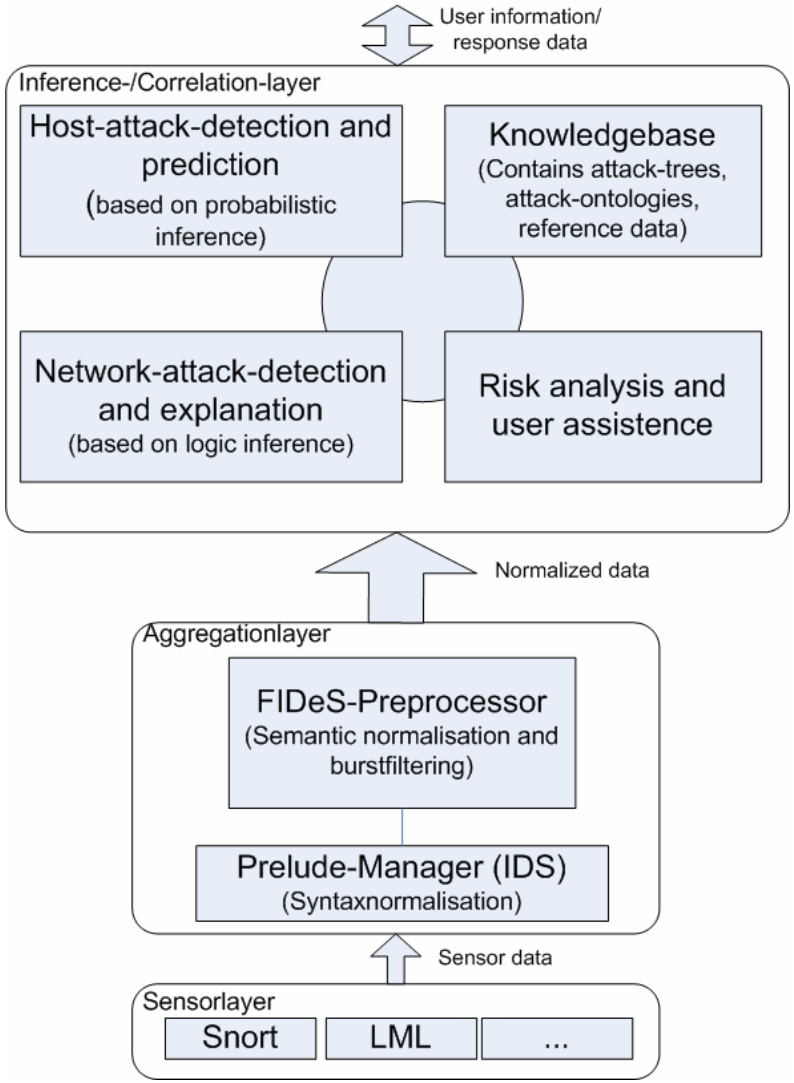
- Auswirkungen von falsch positiven Alarmen
  - **Zu niedrige Schwellenwerte** erzeugen zu viele Ereignisse
  - Mehr Ereignisse als notwendig müssen verarbeitet werden
    - Aufwand & Kosten
  - Zu viele irrelevante Ereignisse verringern die **Verständlichkeit**
  - Zu viele Ereignisse erhöhen die Chance relevante Ereignisse zu **übersehen**
  - Zu viele Ereignisse begünstigen eine ungewollte **Gewöhnung**, steigende **Frustration** und unnötige **Überlastung** beim Menschen
- Auswirkungen von falsch negativen Alarmen
  - **Zu hohe Schwellenwerte** erzeugen zu wenig Ereignisse
  - Relevante Ereignisse werden **„verschluckt“**
  - Eine **Unterforderung** kann die Aufmerksamkeit senken

Projektziele

- Entwicklung eines **intelligenten Assistenz- und Dialogsystems** zum Erkennen von (professionellen) Angriffen auf Unternehmen aus dem Internet
- Behandlung von weit verbreiteten Internetprotokollen wie z.B. HTTP oder SMTP
- Aber auch: Behandlung von **neueren/weiteren Protokollen**
  - BGP – Border Gateway Protocol (zwischen Routern)
  - SOAP: auch missbräuchliche Zugriffe in sicherheitskritischen, IT-gestützten Geschäftsprozessen
  - Perspektivisch: Telekommunikationsprotokolle wie z.B. UMTS, VoIP-Protokolle
- Zusammenführung **verschiedener Datenquellen**:
  - Log-Dateien an Firewalls, auf Betriebssystemebene, auf Applikationsebene

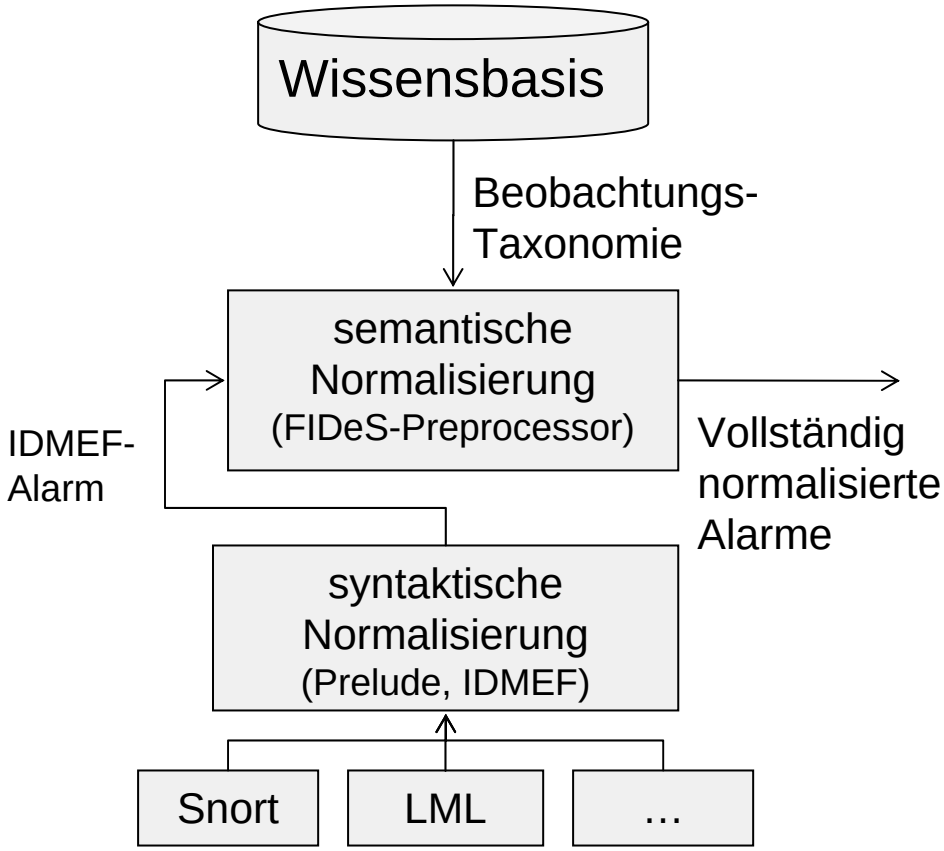
Projektziele und Architektur

- Bereitstellung von plausiblen, nachvollziehbaren **Erklärungen** für einen Angriff
- **Deklarative Beschreibung** von Angriffswissen, Systemen, Systemkomponenten und Gegenmaßnahmen
- **Interaktive Assistenz** bei der Durchführung und Auswahl von Gegenmaßnahmen
- **Vorhersage** über den weiteren Angriffsverlauf
- Bereitstellung einer **Wissensdatenbank** mit Beschreibungen über Angriffe und Gegenmaßnahmen
- **Angriffserkennung** durch Maschinelles Lernen



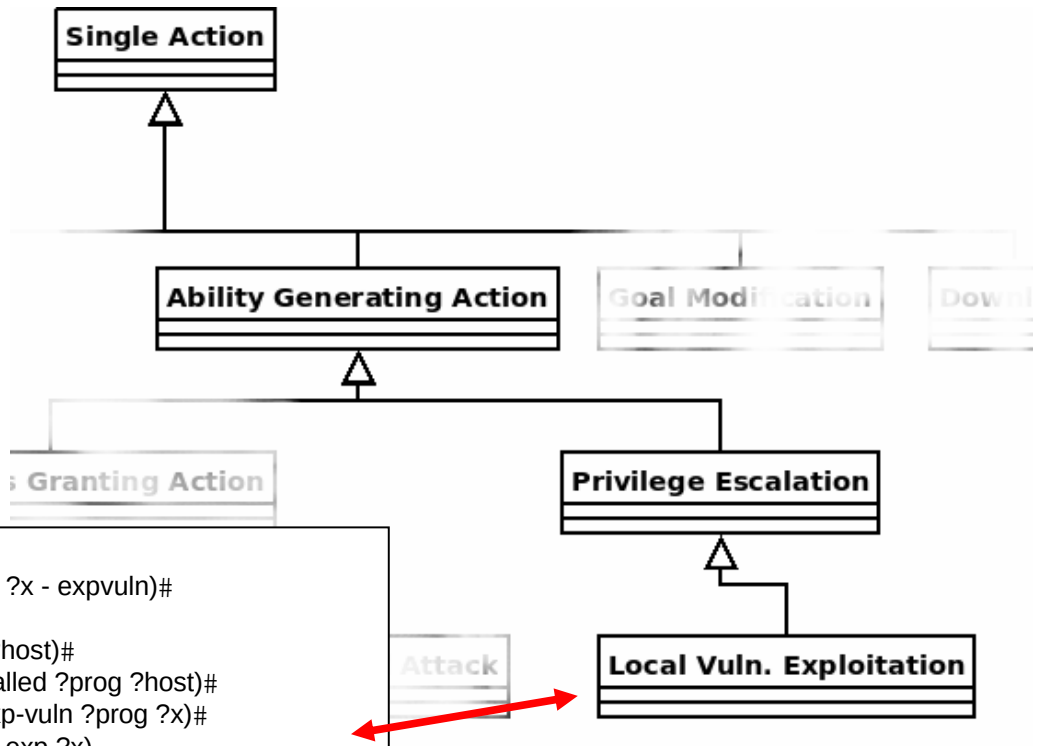
FIDeS Komponenten - Vorverarbeitung

- Ursprüngliche Alarme der Sensoren werden durch **PreludeIDS** gesammelt
- Prelude schickt **IDMEF** Nachricht an Vorverarbeitung
- **Semantische Normalisierung** und „**Burstfiltering**“ durch die Vorverarbeitung
- Normalisierungs-Informationen sind in der **Wissensbasis** hinterlegt



Wissensbasis

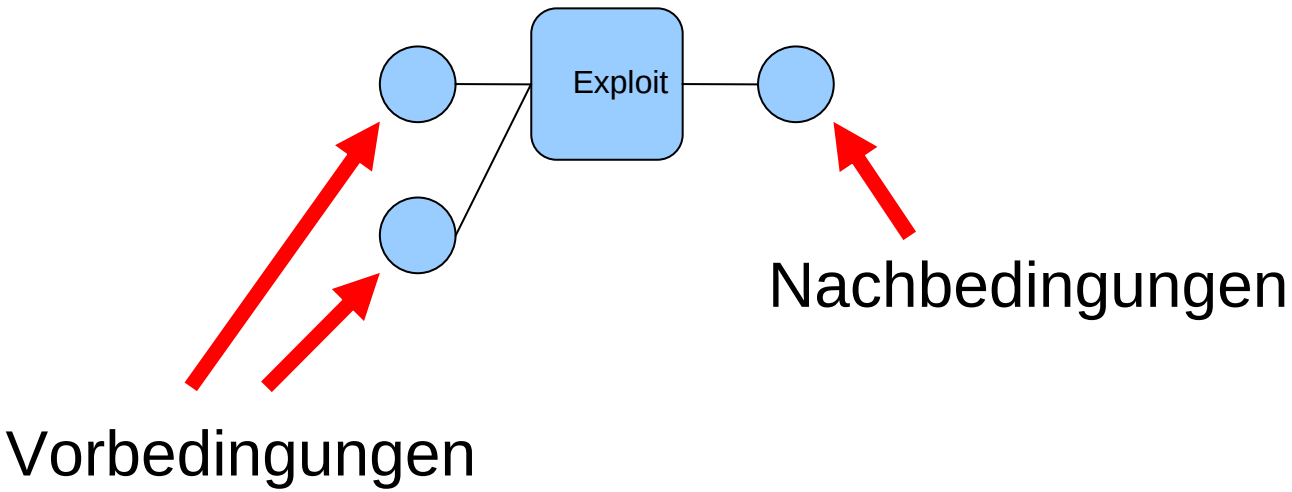
- Taxonomien für Alarme und Angriffsschritte
- Teilpläne und Planungsoperatoren
- Attack-Trees
- …



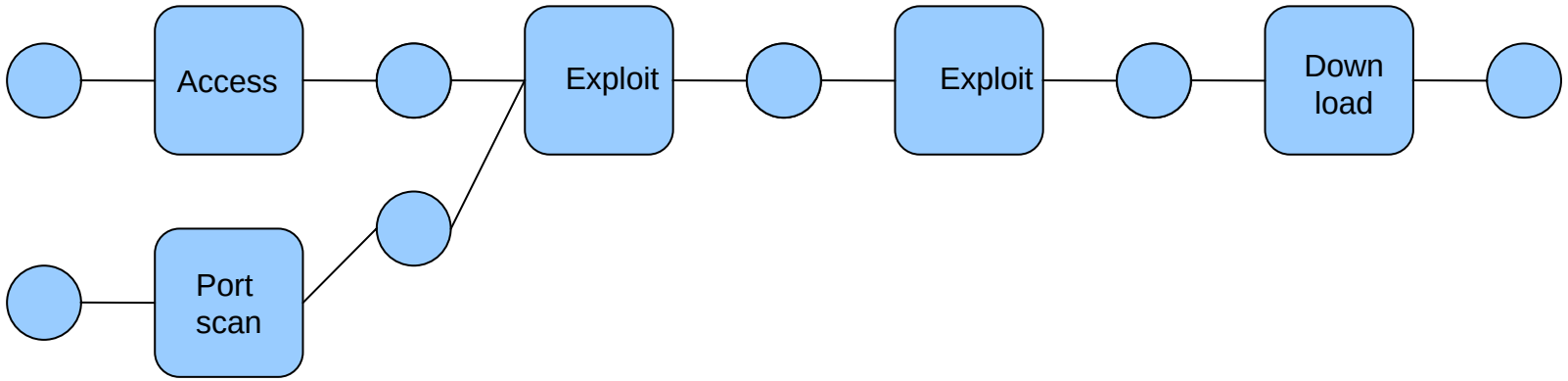

```
(:action exploit-local
        :parameters (?host - host ?prog - program ?x - expvuln)#
        :precondition (
                and         (a-is-on ?host)#
                            (p-is-installed ?prog ?host)#
                            (p-has-exp-vuln ?prog ?x)#
                            (a-knows-exp ?x)
                )#
                :effect (
                and         (a-root ?host)#
                )#
)#
```

Logische Repräsentation von Angriffen

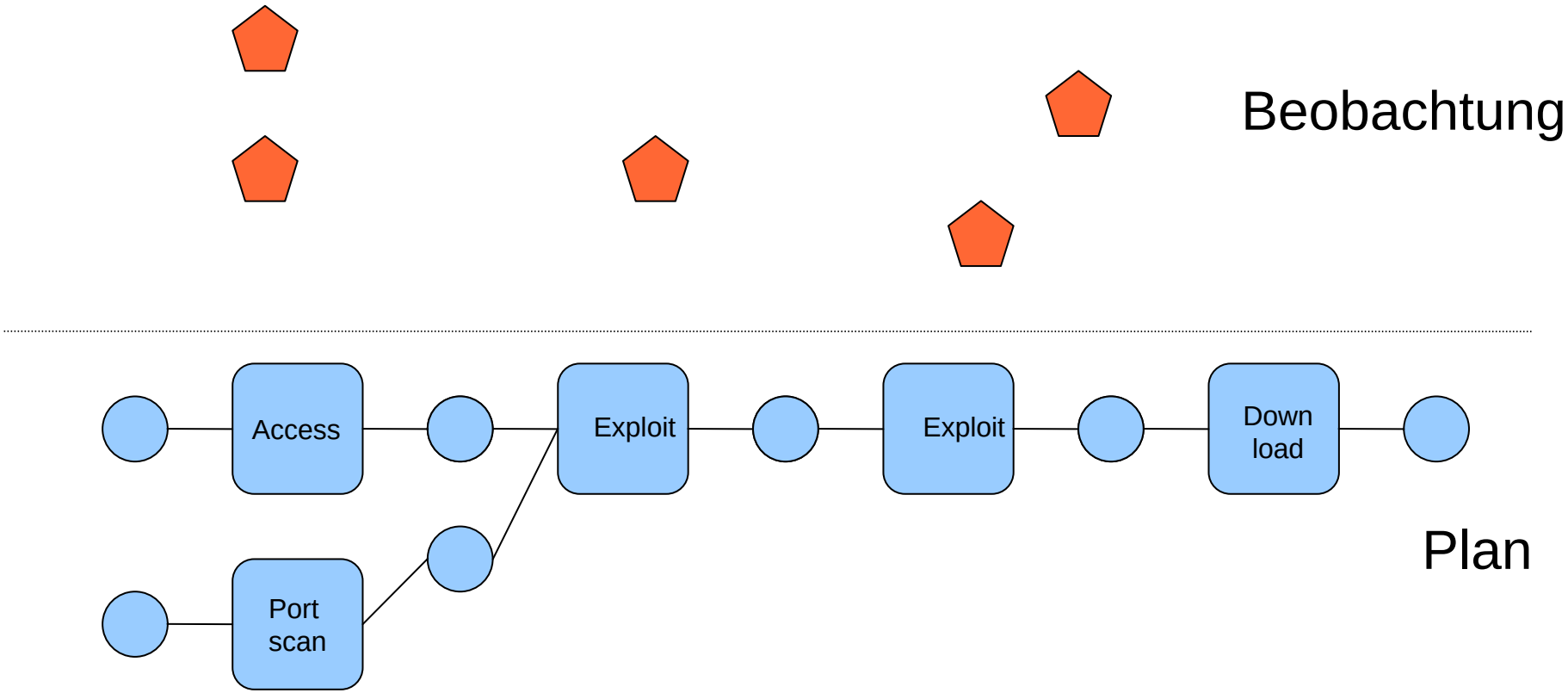
- Ein **Beispiel** zur Veranschaulichung
- Modellierung der Abhängigkeiten von Angriffsschritten durch Experten
- Generieren von Plänen durch dieses Wissen



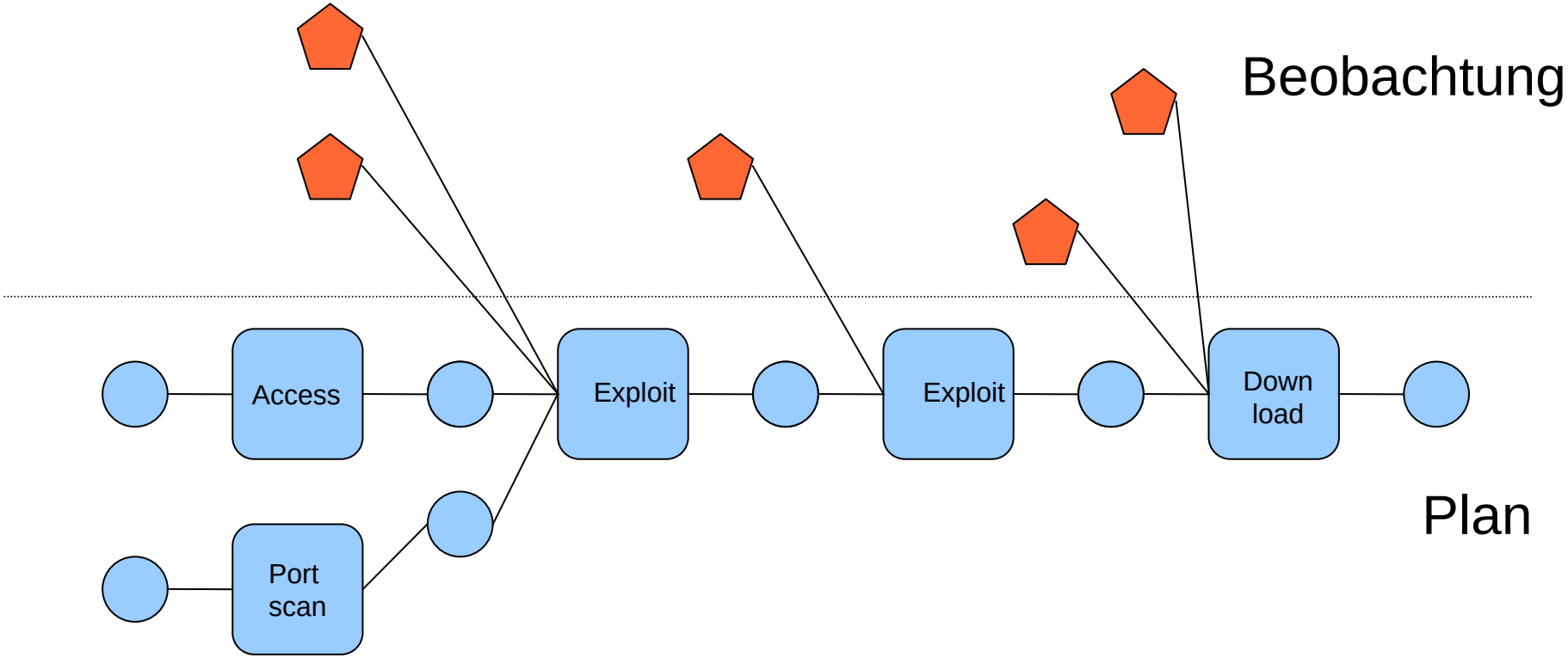
Logische Repräsentation von Angriffen



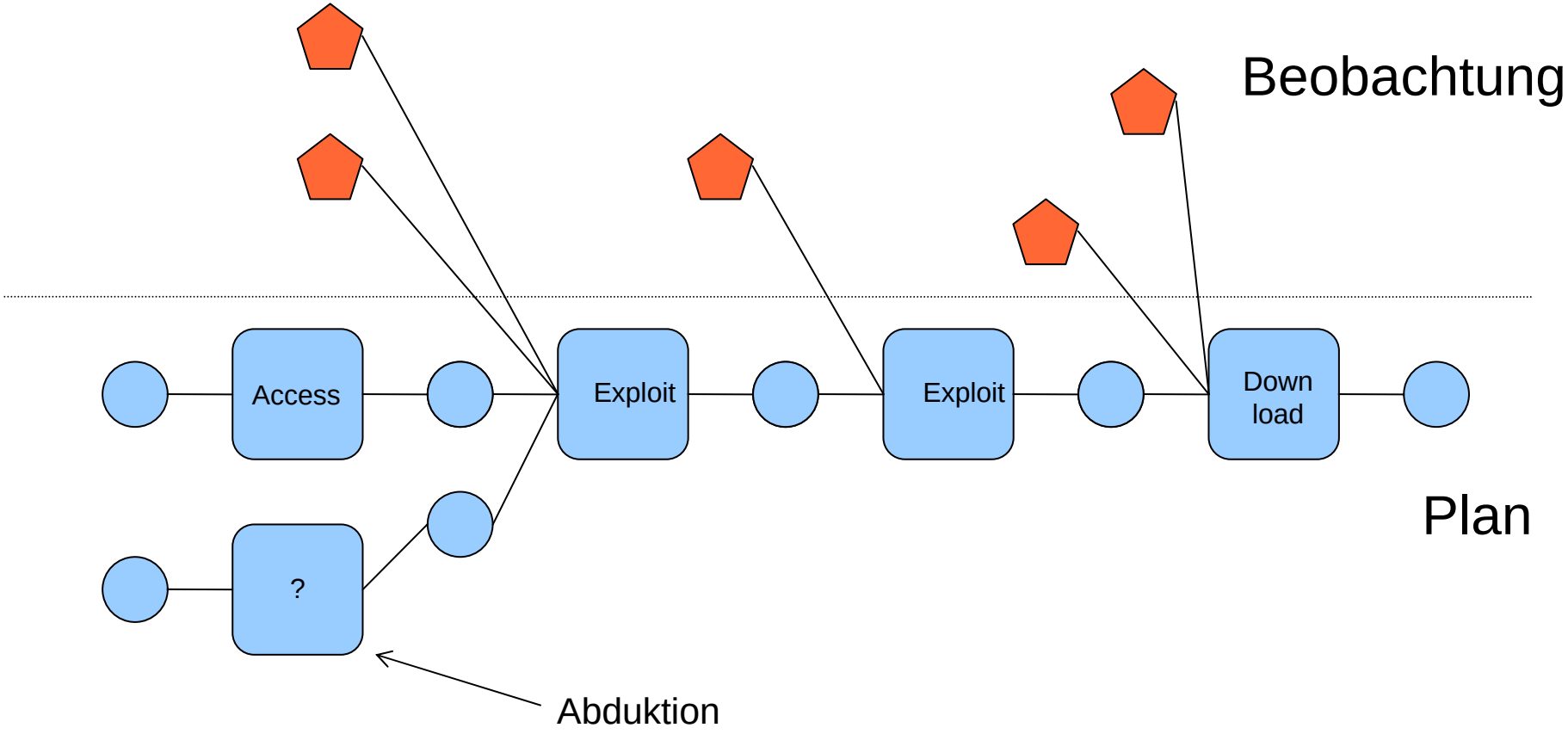
Logische Repräsentation von Angriffen



Logische Repräsentation von Angriffen

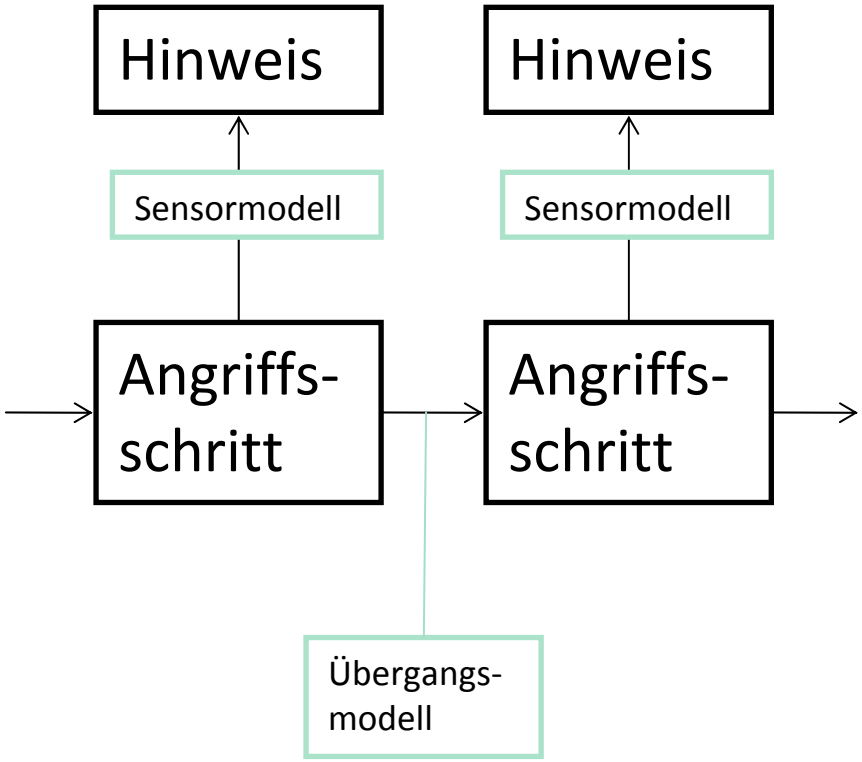


Logische Repräsentation von Angriffen



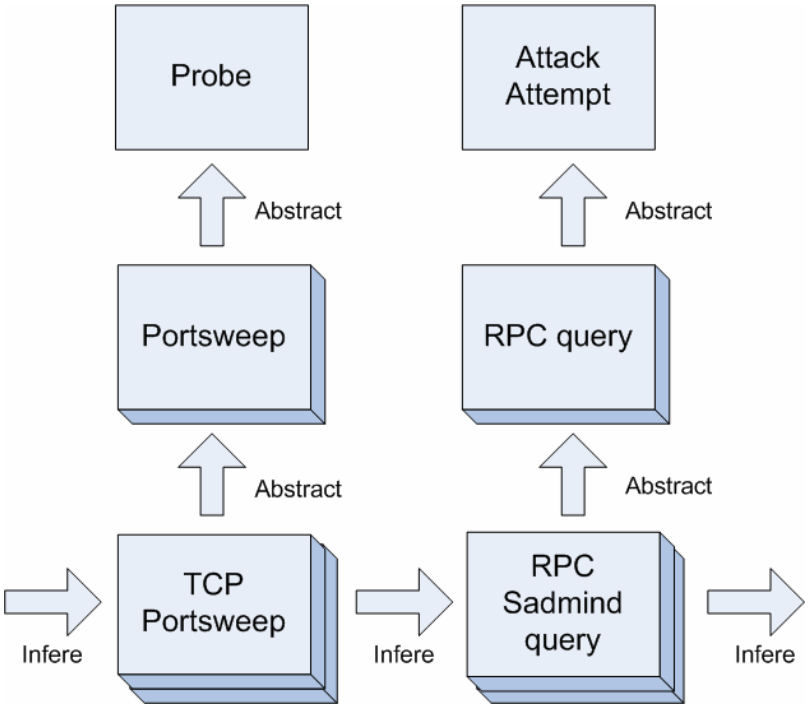
Probabilistische Repräsentation

- Annahmen:
  - Unbekannte Angriffssequenz
  - Alarmsequenz kann Fehler beinhalten
  - Suche nach der wahrscheinlichsten Angriffssequenz für die beobachteten Alarme
- Hidden Markov Model (HMM)
  - Angriffsschritte sind unsichtbar
  - Alarme geben Hinweise auf die Angriffsschritte



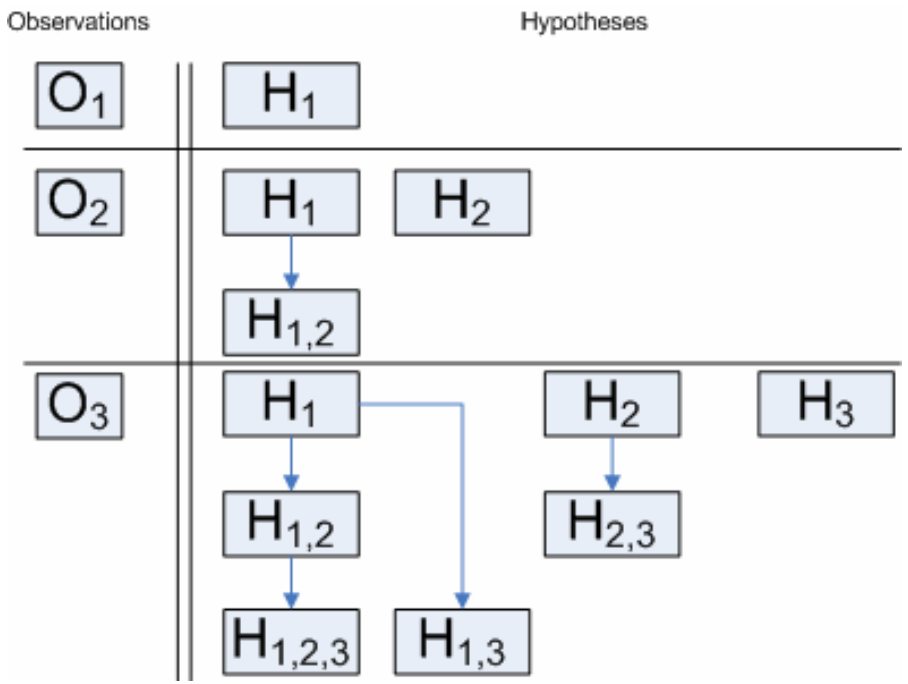
Probabilistische Repräsentation

- Probleme:
  - Wenig Referenzdaten
  - Unsicherheit über genauen Angriffsverlauf, Sicherheit bzgl. ungefährem Angriffsverlauf
- Relationales HMM
  - Ähnliche Beobachtungen basieren auf ähnlichen Angriffsschritten



Langzeit Angriffe

- Problem:
  - Nur wenige Alarme sind auf echte Angriffe zurückzuführen
  - Angriffe erstrecken sich ggf. über einen langen Zeitraum
- Mögliche Lösung:
  - Eine Hypothese für jede Alarmkombination
  - Beschränkung der Hypothesen auf die Wahrscheinlichsten



Gestaltung der Nutzer-Schnittstelle

- Vereinfachung der Wissens-Modellierung
- Experten und Nicht-Experten die Möglichkeiten geben, ihr Wissen in die **Wissensbasis** zu übertragen
- Dialog mit und **Feedback** von dem System
  - KI-Methoden, um **Hypothesen** zu generieren
  - Interaktion mit den KI-Methoden: **Bewertung** von Hypothesen
  - **Trainieren** des Systems (Supervised Learning)
  - **Erheben** was vom Nutzer gewünscht wird und was für ihn interessant ist
- Leichtere **Anpassung** an ungewöhnliche oder neue Datenquellen

Zusammenfassung

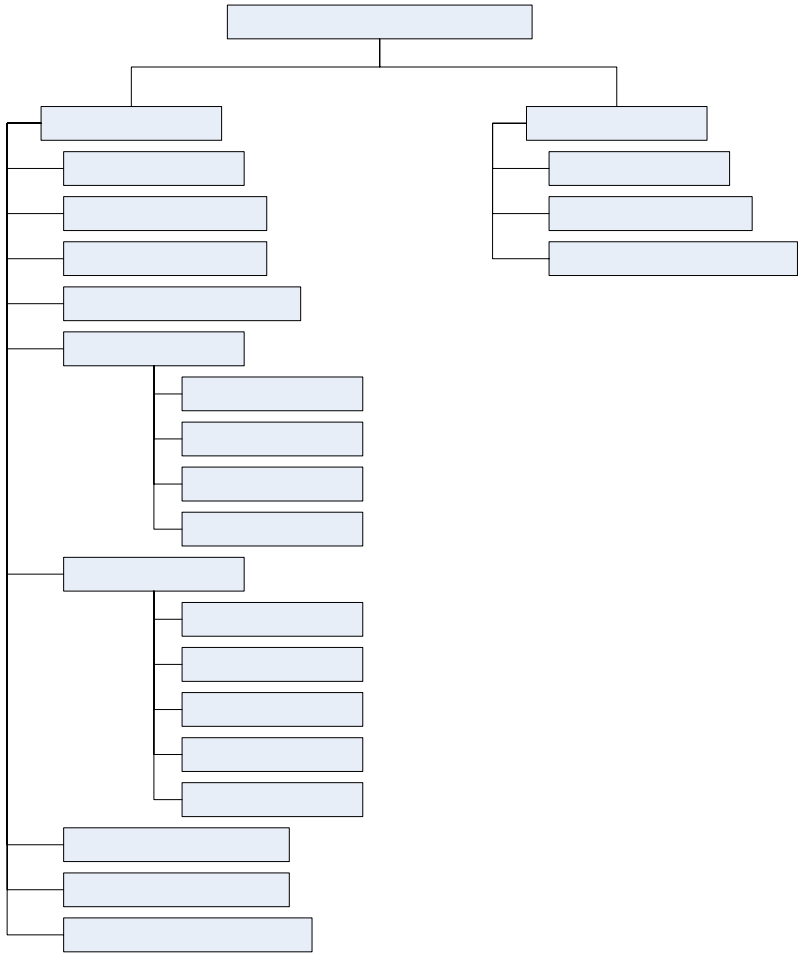
- Intelligente **Korrelation** auf hoher Ebene
- **Erkennung** von (Langzeit)Angriffen
- **Erklärungsgenerierung** und Risikoanalyse durch modelliertes Wissen
- **Vorhersage** von wahrscheinlich folgenden Angriffsschritten
- Vorschläge für **Gegenmaßnahmen**
- **Grundlagenforschung** in der Anwendung von KI-Methoden im Bereich der Intrusion Detection
- **Interaktion und Assistenz** als Grundkonzept

Ende

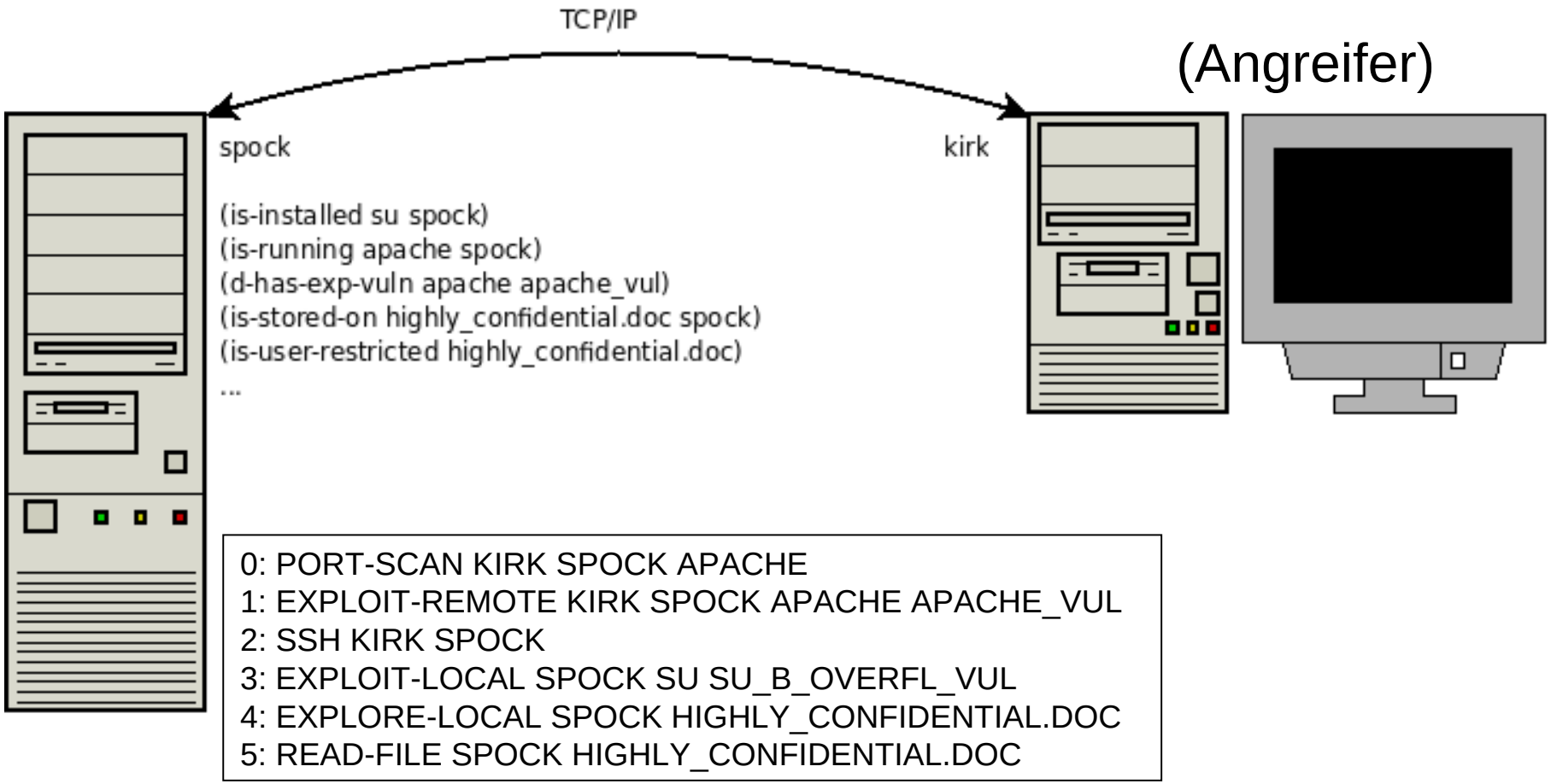
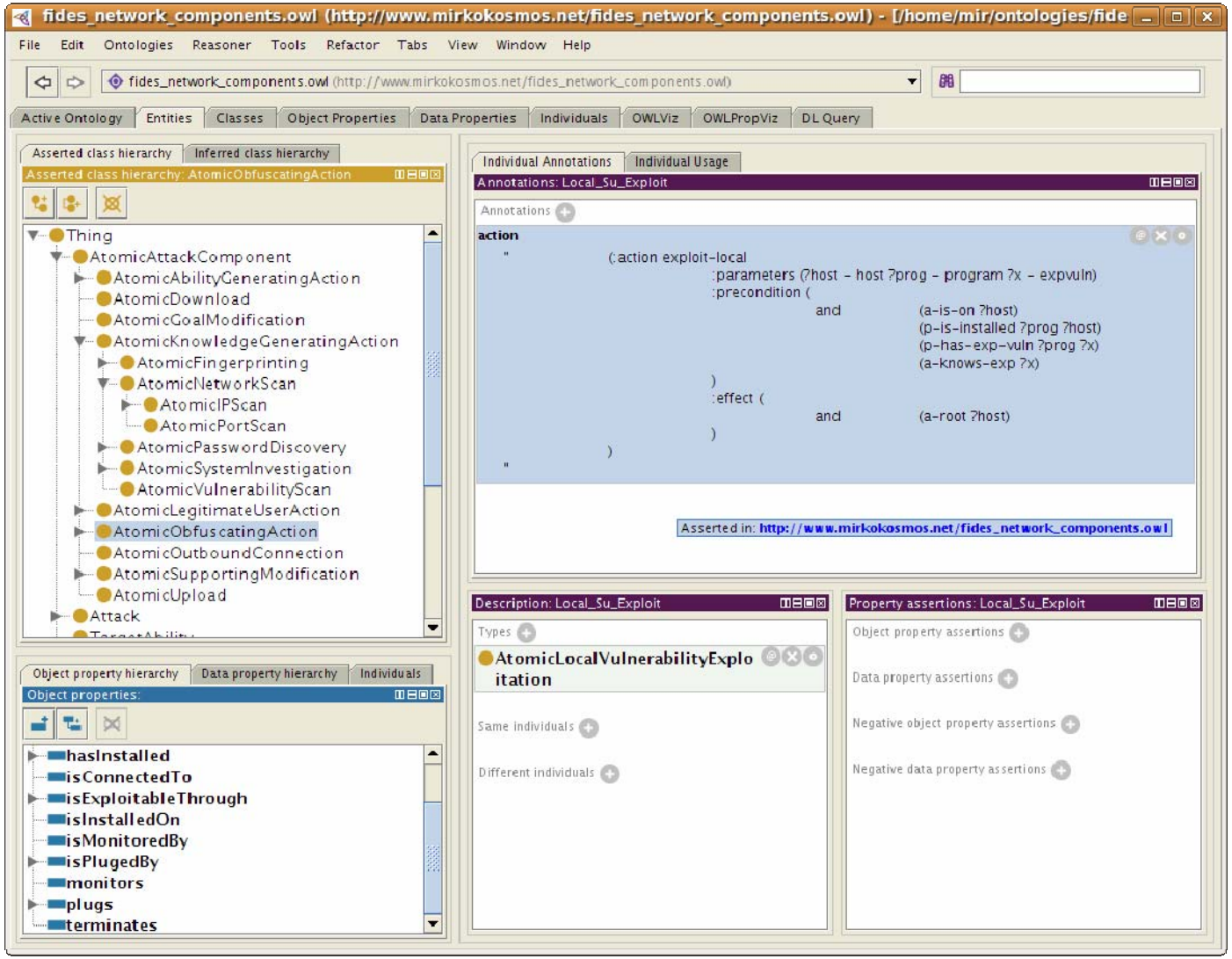
Vielen Dank!

IDMEF

- Prelude zur Sammlung von Ereignissen im System
  - Security Information Management System
  - Schnittstelle zur Anbindung von Sensoren und weiterverarbeitenden Modulen
  - Nutzt das Intrusion Detection Message Format (IDMEF) zum Datenaustausch
- Anbindung an SNORT, Prelude-LML, IAS der FH Gelsenkirchen



Beispiel: Netz und Angriffsplan



Ontologie und Planungsdomäne